IUPAC Division of Organic Chemistry

Organic Nomenclature in the 21st Century

Summary Report of the Working Group
March 1999

Introduction

The object of this project was to establish future needs for organic chemical nomenclature, by means of a survey of nomenclature users.

Procedure and results

During the autumn of 1998 we distributed a request for comment on future nomenclature requirements to suitable internet discussion groups, nomenclature committees, journal editors and ChemWeb members, and also published the request as a letter in chemistry news journals.

During the period October 1998 to January 1999, 139 replies were received, from respondents in 39 different countries. The opinions expressed showed a considerable degree of convergence: the requirements specified by the respondents are summarised in Table 1, which also shows the numbers of people supporting each requirement (more specialised requirements were passed directly to the appropriate IUPAC Commissions). The geographical distribution of respondents is shown in Table 2.

We met in Montreal in February 1999 to consider how the requests in Table 1 might be addressed, and formulated the following recommendations.

Recommendations for IUPAC Division of Organic Chemistry

As the new millenium approaches, fundamental changes are being made to IUPAC`s structure and working procedures and it is appropriate to consider carefully how the Union`s goals and procedures for nomenclature codification and development can be suitably accommodated in this context. In 1992, the Chairman of the Commission on Organic Nomenclature (Henri Favre) wrote `With the emergence of computerized nomenclature, the role of CNOC is obvious. The preferred name should be determined as soon as possible and a strong capacity for using the computer must be developed`. This statement is as appropriate to-day as it was in 1992. Work on nomenclature standards has always been a core activity for IUPAC, and it has always been the ultimate intention to develop systems enabling every molecular structure to be assigned an unambiguous unique name. In organic chemistry, this aim is now close to realisation, with the completion of the `preferred` names project. The availability of a IUPAC-preferred name will facilitate projects not previously considered practical. At the same time, it has become clear that the power of the computer can be harnessed to provide nomenclature tools, and the emergence of world-wide computer networks gives enormous potential for sharing of information and facilities. In this context, we will consider the needs identified by the results of this survey, and make suggestions as to how they might be addressed.

As will become clear, there is much work to be done, and we consider that the requirements could be best addressed and the work coordinated by a small Nomenclature Systems Standing Committee. It will be particularly important to maintain, extend and service the complex and extensive set of interlocking recommendations, developed over many years, that yields the IUPAC-preferred names. The period immediately following publication of these rules will be a critical one: we should expect

Analysis of responses to survey, October 1998-January 1999

Table 1. General requirements

No.	Requirement	Supporters
1.0	*Computerised naming facilities*	
1.1	Generally (freely) accessible computerised structure-to-name conversion via the World Wide Web	22
1.2	Computerised name-to-structure conversion	6
2.0	*Correlation with other systems*	
2.1	Convergence/reconciliation of IUPAC, CAS and other systems and computerised interconversion	14
2.2	Correlation between IUPAC preferred names and CAS registry nos.	5
2.3	Correlation of International Non-Proprietary Names with IUPAC preferred names	1
2.4	Computerised trivial-to-systematic name translation facility	14
3.0	*New databases*	
3.1	Database of trivial (common) names, with translation to structures and systematic names	11
3.2	Database of common biochemical names with systematic equivalents	2
3.3	Database of systematic names for common chemicals	2
3.4	Database of trivial names for “important” compounds	1
3.5	Glossary of names of environmental contaminants	1
3.6	Database of generic names of drugs and adjuvants linked to systematic names	1
4.0	*General rules revision*	
4.1	Simplification of rules	9
4.2	Recognition of more trivial names for complex structures (with translation to systematic)	6
4.3	Establishment of more trivially named parent systems (e.g. polycyclics) for systematic manipulation	10
4.2	Phasing out of trivial names (but see 2.4, 3.1, 3.4, 4.2)	6
5.0	*New standards*	
5.1	Platform-independent standard for chemical structure files/data standard for chemical identity	8
5.2	Universal chemical registry system (avoiding ambiguities of CAS)	2
6.0	*Terminology development*	
6.1	Class Names based on topology	2

Table 2. Geographical breakdown of responses
(39 countries)

USA	46
UK	14
Germany	7
Belgium	5
Canada	5
India	5
France	4
Russia	4
Unknown	4
Ireland	3
Italy	3
Argentina	2
Czech Republic	2
Israel	2
Japan	2
Netherlands	2
New Zealand	2
Portugal	2
Sweden	2
Switzerland	2
Turkey	2
Australia	1
Brazil	1
Bulgaria	1
China	1
Croatia	1
Finland	1
Greece	1
Hungary	1
Iran	1
Jordan	1
Lithuania	1
Macedonia	1
Norway	1
Romania	1
Saudi Arabia	1
Slovak Republic	1
Spain	1
Thailand	1
Venezuela	1

that problems in application will emerge and modifications to procedures will be required for some time. For this purpose, and for continuing development, the members of the Standing Committee will need to possess considerable expertise and experience in nomenclature matters, and their appointment for extended periods should be envisaged. In the development of nomenclature, continuity is essential, to avoid ill-informed disturbances to the system resulting in conflicts amongst its component rules.

There are two themes that emerge strongly from consideration of the responses to the survey:

- the need for computerised naming facilities to be freely available via the World Wide Web
- the need for correlation of IUPAC-preferred names with other names in common use

General requirements (Table 1)

1.0 *Computerised naming facilities*

There are several commercially produced software packages that offer computerised structure-to-name conversion. These include ACD/Name, produced by Advanced Chemistry Development, Autonom, produced by the Beilstein Institute and Nomenclator, by ChemInnovation. We suggest that IUPAC, through its Nomenclature Systems Standing Committee, should maintain close contacts with all such organizations, to ensure that the programs provide current IUPAC-preferred names. Furthermore, we recommend that negotiations take place for IUPAC to acquire a cut-down version of naming software, providing IUPAC-preferred names for the less complex structures, to be made freely available over the Web as an interactive facility from the IUPAC Website.

Name-to-structure conversion programs are not so far advanced. Again the Standing Committee should maintain contacts with organisations providing such facilities, to ensure as far as possible that IUPAC-preferred names can be handled.

Computer algorithms for naming complex inorganic structures are yet to be developed.

2.0 *Correlation with other systems*

It is in the interest of the whole chemistry community to establish links between different names (both trivial and systematic) and codifications for the same structure. Convergence of naming systems is desirable and to be encouraged, but complete convergence (e.g. adoption of IUPAC-preferred names by Chemical Abstracts Service) is very unlikely. Therefore steps should be taken to provide computerised translation facilities. ACD/Name has facilities for providing both IUPAC and CAS names according to the user`s prefence; it would be helpful any cut-down version of ACD/Name provided from the IUPAC Website were to include facilities for conversion of IUPAC to CAS names and vice versa. Other correlations could best be approached via databases of limited content (see 3.0).

3.0 *New databases*

Responses from the survey point to the need for various databases providing structures and both systematic and trivial names for limited lists of compounds. We suggest that IUPAC carries out a survey of available data collections of this type, and evaluates the feasibility of incorporating IUPAC-preferred names and making the results generally available over the Internet. Some such lists are quite small (e.g. the BSI Names for Chemicals used in Industry); some are very large (e.g. the Merck Index, the European Inventory, and the Dictionary of Chemical Names and Synonyms

from Synapse Information Resources). It would be possible to compile selective data collections from various sources, giving structure, IUPAC-preferred name, CAS name and registry number, and common trivial names (available from the CAS Common Names Handbook). Such projects would be valuable but labour-intensive, and we consider that they could not be carried out satisfactorily unless on a commercial basis, with funds made available to the Nomenclature Standing Committee

We emphasise the importance of ensuring that publications purporting to quote IUPAC names use the IUPAC-preferred name in future. For example generic names for pharmaceuticals are published with IUPAC equivalents in various contexts such as national pharmacopoeias; the bodies responsible should be alerted as soon as IUPAC-preferred names are available.

In this context IUPAC should consider seriously the possibility of setting up a naming service at the Secretariat, as soon as the principles for selecting IUPAC-preferred names are firmly established. The service could be either free or cost-recovered, and would need at least one permanent member of staff with access to full versions of commercial naming software. The staff would be able to use the available software to generate names in most cases, but would also be expected to deal with cases not handled satisfactorily by the software, and would have the possibility of referring particularly difficult problems to designated IUPAC `consultants` (e.g. former members of Commissions). Scientific responsibility for the service would rest with the Nomenclature Systems Standing Committee.

4.0 *General rules revision*

Several survey responses asked for simplification of naming procedures. It seems to us that this will not be needed when computerised naming is widely available. Only experts will need to appreciate the full ramifications of nomenclature rules. However, there is scope for making names of complex structures simpler, by assigning `trivial` names to a greater number of large molecules as parent systems for systematic modification. We recommend that this possibility should be explored in particular for polycyclic systems. The planned further development of phane nomenclature will also alleviate the problem.

We do not agree with those respondents who recommended that the use of trivial names should be phased out or otherwise deprecated. The number of trivial names that will be IUPAC-preferred is small. However, there is a need to maintain links with literature in which obsolescent trivial names are used, and this can be addressed by data compilations of the kinds referred to in section 3.0. Furthermore, there is a continuing need for trivial names to simplify current literature, provided that the IUPAC-preferred equivalent can be readily ascertained.

5.0 *New standards*

A few years ago. IUPAC`s Committee on Chemical Databases, now absorbed into the Committee on Printed and Electronic Publications (CPEP) considered adopting a standard IUPAC chemical structure file format. However, for a variety of reasons, a proposal document was never received, and the project lapsed. The requests for such a standard in the present survey lead us to recommend that the project should be resurrected, and that IUPAC should evaluate the various formats currently in use.

As an extension to this project, it would be appropriate to develop an electronic `chemical identity` standard, for use in data transmission. This might be a structure in the standard format or perhaps a registry code. One of the respondents to the survey pointed out the unsuitability of the CAS registry

system for this purpose (many cases of multiple registry number for a single structure), and we recommend discussion with CAS about possible ways forward (perhaps a IUPAC-preferred registry number).

6.0 *Terminology development*

Requests were received for terminology based on topological considerations. To us this implies extension of the Class Names glossary to include more terms referring to molecular shape in the rapidly developing areas of macromolecular (particularly macrocyclic) chemistry. The recent literature should be searched for new terms that would benefit from evaluation and subsequent recognition by IUPAC.

Conclusion

On the basis of the foregoing considerations, we strongly recommend that responsibility for future maintenance and development of naming systems be placed in the hands of a Nomenclature Systems Standing Committee. We believe that it would be preferable for this Committee to cover nomenclature systems for organic, inorganic and macromolecular compounds. It would not be expected to deal with definitions of terms (general chemical terminology).

Jean-Claude Richer
Alan D McNaught